

TRAIT MATH ANXIETY IN HIGH SCHOOL STUDENTS AS ASSESSED BY ELECTROPHYSIOLOGICAL MONITORING DURING A STANDARDISED MATH TEST

DSS SHIVARANJANI, Assistant Professor, sivaranjani1238@gmail.com, Rishi UBR Women's College, Kukatpally, Hyderabad 500085

ABSTRACT

Math anxiety (MA), i.e. a trait factor describing the feelings of tension, apprehension, and fear during mathematics-related situations, has attracted increasing interest in recent years, due to its importance in people's daily life and career development, especially in our modern digital world. Although the measurement of individuals' trait MA has mostly relied on self-reported psychological scales, emerging studies are seeking for objective measurement by using behavioral or neurophysiological data. The present study, for the first time, investigated the neurophysiological signatures of trait MA in a cohort of high school students during their 90-minute real final-term math exam. Wrist-worn wearable devices were used for recording their autonomic nervous system activities, including skin conductance (SC) and heart rate (HR). The calculation of pairwise correlation revealed that both SC and HR could reflect the individual's math evaluation anxiety (MEA) score, which is one of the two sub-components of trait MA. Specifically, the tonic level of SC was negatively correlated with MEA during the 5-minute pre-exam period when the students were anticipating the exam, whereas HR was positively correlated with MEA at two later time windows during the exam (63-65 minutes and 82-85 minutes, after the start of the exam). A leave-one-out regression analysis revealed a correlation ($r = .349$, $p = .094$) between the self-reported MEA scores and the scores predicted by these neurophysiological signatures. Our findings provide neurophysiological evidence of trait MA in a real-life context and demonstrate the potential of implementing an objective measurement of trait MA based on neurophysiological signals.

I. INTRODUCTION

Math anxiety (MA) is considered to be a trait-level disposition [1], which describes the general feelings of tension, apprehension, and fear when dealing with a wide range of mathematics-related tasks and situations [2], [3]. MA is a kind of anxiety that is independent of general anxiety or test anxiety [3]–[5]. An increasing number of students worldwide suffer from MA [6], which leads to their negative attitude

towards math, failure in math courses, the avoidance of learning and using math or numeric contents even in their future work, etc. [3], [7], [8]. Negative correlations between MA and math achievement have been reported across multiple developmental stages [3], [6], [9], [10].

The measurement of trait MA is mainly based on self-reported scales that require individuals to read MA-related statements and make judgment about their degree of

agreement. The most widely used scales include the Mathematics Anxiety Rating Scale (MARS-A/E) [11], [12], and Math Anxiety Scale for Children (MASC) [13]. While these self-reported scales have the advantages of straightforwardness and cost-effectiveness, they are prone to recall and social desirability biases. For example, people may underestimate their feelings during math-related situations when they fill up the scales afterwards, and they could also deliberately fake his/her responses for a variety of reasons. These issues limit the effectiveness of the method for MA measurement. Implicit methods like the arithmetic-affective priming task [14], which measures the MA by how the affective stimuli influence the reaction time to the arithmetic questions, have been proposed to address these problems. However, these measurements need participants to perform tasks and therefore cannot be used for daily monitoring of MA, especially for high school students who suffered a lot in math-demanding learning and tests.

Studying the neurophysiological basis of MA may provide a way to solve this issue, as the neurophysiological signals could provide an objective description of people's reaction during math-related situations. Efforts in this direction date back to at least 1984, when Dew et al. [15] examined the representation of MA in the autonomic nervous system (ANS) and reported significant correlation between self-reported MA and ANS responses of skin conductance (SC) level and heart rate (HR) in a test-like situation. Subsequently, Faust [16] reported significantly higher HR for a group of individuals with higher MA compared to a

group of individuals with lower MA when dealing with difficult math tasks. However, recent studies have mainly focused on the central nervous system (CNS) [17]. For instance, individuals with higher MA have been shown to exhibit more fear-related activation in the right basolateral amygdala [18], reduced frontocentral and centroparietal processing of numeric information [19], more diffused and unstructured functional network related to corrupted working memory process [20], as well as stronger gamma band activity relating to greater attentional bias towards negative emotion during arithmetic problem solving [21]. Although the above-mentioned studies are mostly focused on the math task performing stage, researchers are beginning to pay attention to the anticipation period. For instance, anticipation of a math task has been reported to elicit activation in brain regions associated with pain perception for people with higher MA [22]. Stronger beta band oscillation and P300 amplitude associated with increasing attentional resources usage [21] as well as more effective brain functional organization in their attempt to regulate negative emotions [23] were also found in individuals with higher MA when anticipating upcoming math tasks.

The above-mentioned neurophysiological findings, however, may be undermined by the ecological validity of the experimental paradigms. Most of the studies have used simple arithmetic tasks, such as numeric comparison, addition, division and multiplication, to elicit MA in laboratory environments [19], [21], [23]. These laboratory-based paradigms, however, do

not resemble real-life conditions in terms of task complexity, duration, consequences of failure, etc. As a result, the participants, most of whom are college students, might not sufficiently engage in these math tasks. Such a concern calls for the introduction of new paradigms that could provide a better approximation of real-life MA. Although MA-related neurophysiological studies have not adequately addressed the ecological validity issue, there are good examples in the fields of general anxiety and education. For instance, classical general anxiety tasks, such as public speaking and job interviews could mimic real-life anxiety scenarios very well [24]–[26]. The recent education related studies have moved even further into real-world classroom environments, by recording neurophysiological signals while the students are engaged in their actual classes [27]–[29]. Researchers believe that our nervous system may have a stronger ‘tuning’ to naturalistic tasks than laboratory-based ones, and examining neural processing under dynamic and natural conditions may facilitate a deeper understanding of adaptive human neural mechanisms [30], [31].

Despite these potential advantages, the movement towards high ecologically valid scenarios poses challenges to neurophysiological recording techniques. Traditional methods for monitoring CNS activities, like functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), are not suitable due to factors, such as immobility, interference with normal activities, etc., although there are recent promising developments and preliminary applications, as reviewed by Siddharth et al. [32]. In

contrast, ANS signals, such as HR and SC, which can be acquired with minimal influence on the normal activities of the participants (e.g. wristband-like or chestband-like devices), appear to be the most feasible neurophysiological approach available. Besides the above-mentioned ANS studies on MA, there is ample evidence on the correlations between ANS activities and general anxiety. For example, increases in HR and SC level (SCL) are often observed during anxiety-inducing tasks, such as stress interview [33], and public speech [24], [34]. Although studies on MA-specific anticipation are lacking, the ANS representation of other kinds of anticipation has been frequently reported, but with inconsistent findings. For instance, increased HR and SCL were reported when the participants were anticipating a threat or future shock [35] or preparing for a speech [34], [36]. Instead, others found HR reduction when participant adults who stutter were anticipating a stressful speech [35]. Furthermore, there are also studies focusing on the neurophysiological signatures of individual differences in anxiety. A positive correlation was reported between the average HR and anxiety level during public speaking [24], and a recent study reported a negative correlation between the average electrodermal activity (EDA) and anxiety level during public speech [26]. These inconsistencies may be explained by the differences in the tasks (anticipating future shock v.s. public speech) and populations (normal people v.s. people who stuttered) [22], [26], [34]–[37]. Indeed, the mappings between physiological signals and psychological status reported in

previous studies have been problematic for cross-task and cross-population generalization [38], [39], which further highlights the importance of ecological validity: if the experiment could be conducted in the typical scenarios and population related to MA, the corresponding findings are more likely to be valid and useful for possible practical applications.

The present study seeks to explore the neurophysiological signature of MA using a highly ecologically valid paradigm. To this end, we measured the SC and the HR in a cohort of high school students using wearable neurophysiological monitoring during their actual final-term math exam. The math exam took 90 minutes and differed substantially from simple math tasks in task complexity. More importantly, the students were expected to be fully engaged, as it was the most important test of the semester. The 90-minute exam period as well as a 5-minute pre-exam period were recorded, allowing us to investigate both the task stage and the anticipation stage. As MA is a trait-level variable, in this study the analysis was performed at a single participant level that focused on individual differences, according to the majority of previous studies on MA. To fully examine individual differences, we adopted a correlational approach to data analysis, rather than dividing the students into groups. A neurophysiological signal would be regarded as MA-related if it was significantly correlated with the individual-based self-reported MA scale scores. To take full advantage of our 95-minute long recordings, both the SC and HR were further decomposed into 1-minute segments before performing the correlational analysis,

allowing us to inspect the temporal dynamics of the neurophysiological activities. Ultimately, a leave-one-out regression analysis was conducted to establish an overall predictive model of trait MA by using these neurophysiological signals.

II. MATERIALS AND METHODS

A. PARTICIPANTS

All participants were from a regular high school in Beijing. 139 students (65 females; age: 15-17) from grade 11 were invited to fill out the MA scale. Among them, 35 students (15 females; age: 15-17) volunteered to wear a wristband for recording their neurophysiological data during their finalterm math exam. The study was conducted in accordance with the Declaration of Helsinki and the protocol was approved by the ethics committee of the Department of Psychology, Tsinghua University. All the participants and their legal guardians gave written informed consent.

B. MATERIALS

1) MATH ANXIETY SCALE

The Chinese version of the MASC [13] was used to measure the self-reported MA of the participants. Two main factors, Mathematics Learning Anxiety (MLA) and Mathematics Evaluation Anxiety (MEA) [12], [40], [41], were extracted from MASC: MLA involves anxiety and negative emotion toward activities and processes of learning mathematics, and MEA involves anxiety and negative emotion associated with being evaluated or tested in mathematics. Together

with the MA total score from the MASC, three scores (MLA, MEA and MA total) were obtained per participant.

2) NEUROPHYSIOLOGICAL DATA ACQUISITION

A custom-designed wristband (Psychorus, HuiXin, Beijing, China) was used to record SC and HR. SC was measured by surface electrodes with conductive gels at a sampling rate of 40 Hz. HR was measured by the photoplethysmography (PPG) method at a sampling rate of 20 Hz. Three-axis acceleration was recorded at 20 Hz as well but not used in the present study.

3) ACADEMIC PERFORMANCES

The final-term math exam scores from all 139 participants were collected. The duration of the final-term math exam was limited to 90 minutes. The exam covered topics of algebra, geometry and calculus taught during the semester and the problem types consisted of multiple-choice, filling-blanks, and comprehensive problem-solving questions, which were far more complicated than the simple laboratory-based math tasks. To control for the specificity of the MA scales, their Chinese exam scores were also collected. Overall, the 139 students scored from 15 to 119 (full marks: 120) for their math exam, and from 36 to 91 (full marks: 120) for their Chinese exam. For the students involved in neurophysiological recordings, their math and Chinese scores varied from 35 to 119, and 53 to 85, respectively.

C. PROCEDURE

The final-term math exam took place on January 16th, 2019. The 35 participants who volunteered for the neurophysiological recordings were asked to wear the wristbands on their left wrist (right wrist for the 2 left-handed students) 10 minutes before the exam with the help of the experimenters (e.g. pasting the conductive gels, etc.). The recordings started 5 minutes before the exam and lasted until the end of the 90-minute exam. In the pre-exam period, the students were sitting quietly waiting for the exam papers to be distributed. According to the requirement of the school, the invigilator announced the exam time at the 60th minute and 80th minute of the exam, respectively.

All 139 participants completed the MASC on January 2nd, 2019, during a psychology course two weeks before their math exam. The scales were collected and analyzed by our experimenters. The students were explicitly informed that their MASC scores would not be revealed to their teachers, classmates, or parents.

D. DATA ANALYSIS

1) DATA PRE-PROCESSING

Given the specificity of the final-term exam, the experimenters were only allowed limited time before the exam to prepare the device, leaving no time to check the signal quality before the start of the recording. Therefore, the signal quality was expected to be lower than laboratory-based studies and it is crucial to apply proper post-hoc artifact rejection procedures. In this study, the major issue was the possible loose contact between the wristband and the skin, which would

result in an unchanged very small conductance (correspondingly a very high impedance) of the measured SC signals. In practice, such a situation will lead to a conductance level at the lower limit (.007 micro Siemens) of the measurement range of the device. Accordingly, the data loss ratio at the individual level, defined as the ratio of the lower limit SC values to the whole recording time, was calculated to evaluate the data quality for each subject. The data loss ratio showed a clear separation of the participants into two groups: 24 out of the 35 participants had a data loss ratio below 10% (average: 0.92%, range: 0.00-8.59%), while the other 11 participants had at least 30% data loss (average: 72.02%, range: 32.18-100.00%). Such an segregation could be due to their device preparation conditions: a tightly-worn device could record high-quality data throughout the exam time, whereas low-quality data might imply the device was loosely worn from the beginning. Therefore, 30% was chosen empirically as the threshold to reject data from the 11 participants whose data loss ratio was >30%. The data from the remaining 24 participants (10 females) were retained for further analysis.

The PPG signals were computed and translated into HRs by a joint sparse spectrum reconstruction algorithm implemented in the HuiXin software package, which is known for its robust performance against daily activity artifacts [42]. The output HR data were organized at a 1 Hz pace. The SC signals were decomposed into the tonic SCL and the transient SC response (SCR) using a continuous decomposition analysis (CDA)

method. Although the SCL mainly shows the low-frequency drifts in the SC data, the SCR reflects the rapid phasic changes [43]. Given the complexity of the math exam task, the integration of SCRs (iSCR) was calculated to represent the overall SCR in a certain time period, in order to avoid possible influences by the usually arbitrary decision of the thresholds for peak detection and event definition [27], [44]. The SC signals were downsampled to 20 Hz and then decomposed into the SCL and SCR. The output SCL and iSCR data were organized at the 20 Hz sampling rate.

Afterwards, both the HR and SC (SCL and iSCR) data were further averaged over non-overlapping 1-minute time windows, resulting in 95 samples per participant for each of the neurophysiological indicators (i.e. HR, SCL, iSCR). The samples influenced by the loose contact issue were rejected, resulting in a total of 55 samples out of the 95 (samples per participant) \times 24 (participant) = 2,280 samples. In other words, the data rejection rate of the 24 retained participants was 2.4%.

2) CORRELATION WITH MATH ANXIETY

A simple Pearson's correlation method was used to investigate the neurophysiological signature Pearson's correlations between the MA scores and each of the 1-minute-based neurophysiological indicators were computed. The time scale of 1 minute was empirically selected, as it provided a reliable neurophysiological measurement (i.e. by averaging over 60 1-sec data points), yet having a fine resolution to describe the temporal dynamics of the MA-related

neurophysiological activities over the 95-minute period. Three types of neurophysiological indicators were used, including HR, SCL and iSCR. Three types of MA scores were used as well, including MEA score, MLA score and MA total score. Pairwise correlations from all possible combination of the neurophysiological indicators and the MA scores were calculated, based on the 1-minute neurophysiological data.

TABLE 1. Correlations between MA scores and academic performance

		MLA score	MEA score	MA total score
Standardized score of Math exam	<i>r</i>	-.328***	-.269**	-.422***
	Sig.	<.001	.001	<.001
Standardized score of Chinese exam	<i>r</i>	-.089	.161	.051
	Sig.	.293	.058	.553

Note. ** $p < .01$, *** $p < .001$

III. RESULTS

A. MATH PERFORMANCE AND MATH ANXIETY

Factors of MLA and MEA were derived from MASC by factor analysis, according to previous studies [12], [40], [41]. MEA and MLA subscales consisted of 9 and 12 items of MASC, respectively. Cronbach's alpha was .945 for all items in MASC, .916 for MEA subscale and .924 for MLA subscale, indicating the good reliability of the MA measurement. As summarized in Table 1, significant negative correlations were observed between the participants' math performance (i.e. final-term math exam scores) and their MA, as reflected by all three types of MA scores (MLA: $r = -.328$, $p < .001$; MEA: $r = -.269$, $p < .01$; MA total: $r = -.422$, $p < .001$). The negative

correlations are in accordance with previous research [3], [54]–[56]. No significant correlation was observed between the Chinese performance and MA, indicating the good specificity of our MA measurement.

To determine whether the volunteers who participated in neurophysiological recordings were a representative sample, the distribution of the MA scores (MLA, MEA and MA total) and academic performance (normalized Math scores and normalized Chinese scores) were summarized for all the 139 students, the 35 students who agreed to participate in the neurophysiological recording, and the 24 students whose data were included in the data analysis, respectively. The distributions of the MA scores of the 3 groups are largely similar (Fig. 1a-1c). The math performance of the 24-student group happened to have a distribution more towards the higher score direction (Fig. 1d), but their performance in Chinese remained similar to that of the 139-student group (Fig. 1e).

B. NEUROPHYSIOLOGICAL SIGNATURES OF MATH ANXIETY

The data in Fig. 2 show the temporal dynamics of the grand average SCL, iSCR and HR during the 95-minute recording time. While there was a continuous increase of the SCL as the exam proceeded (Fig. 2a), the iSCR remained relatively stable with transient peaks from time to time (Fig. 2b). The HR increased rapidly during the 5-minute pre-exam period and kept fluctuating in the subsequent 90-minute exam period (Fig. 2c).

The cluster-based permutation analysis indicated a significant correlation between the neurophysiological recordings and MA, as represented in three clusters (all with permutation $p < .001$) in the recorded data, as shown in Fig. 3. The earliest neurophysiological signature was reflected by SCL at the preexam period (Fig. 3a). The SCL negatively correlated with MEA scores, with an average correlation coefficient close to $-.50$ ($p < .001$), implying that the participants with a higher level of MEA had lower SCL when anticipating the upcoming math exam. The MEA and the SCL were also negatively correlated during the exam period, although not significantly. The HR, however, were positively correlated with MEA scores at a later period (i.e. 63-65 minutes and 82-86 minutes, $p < .001$ for both clusters, Fig. 3c). The higher MEA of participants was associated with a higher HR at these time periods. No significant correlations were found for iSCR (Fig. 3b). In addition, MLA and MA total scores were not significantly correlated with the neurophysiological recordings. The temporal dynamics of the correlation values for MLA is shown in Fig. 4.

Due to the widely acknowledged correlation between MA and math performance [3],

[54]–[56], the neurophysiological differences between higher and lower MEA participants might be explained by their different math performance levels. To rule out the possible influence of math performance, partial correlations between MEA and SCL (or HR) were calculated with controlled math performance (served as a covariate) in the three clusters of interest. As shown in Table 2, the correlations between these neurophysiological signatures and MEA remain at the same magnitude when the math performance was controlled. The scatterplots between SCL (or HR) and MEA at typical time points are shown in Fig. 5.

C. PREDICTION PERFORMANCE OF TRAIT MATH ANXIETY WITH NEUROPHYSIOLOGICAL SIGNATURES

Regression analyses based on the 1-minute-based SCL and HR data revealed a cross-validated correlation between the predicted and self-reported MEA scores at a moderate positive level ($r = .349$, $p = .094$), although not significantly. The scatterplot showing the individual's predicted and self-reported MEA scores is shown in Fig. 6. However, the regression analyses for predicting the MLA and MA total scores, however, only reached weak correlations (MLA: $r = .113$, $p = .600$; MA total: $r = .050$, $p = .815$).

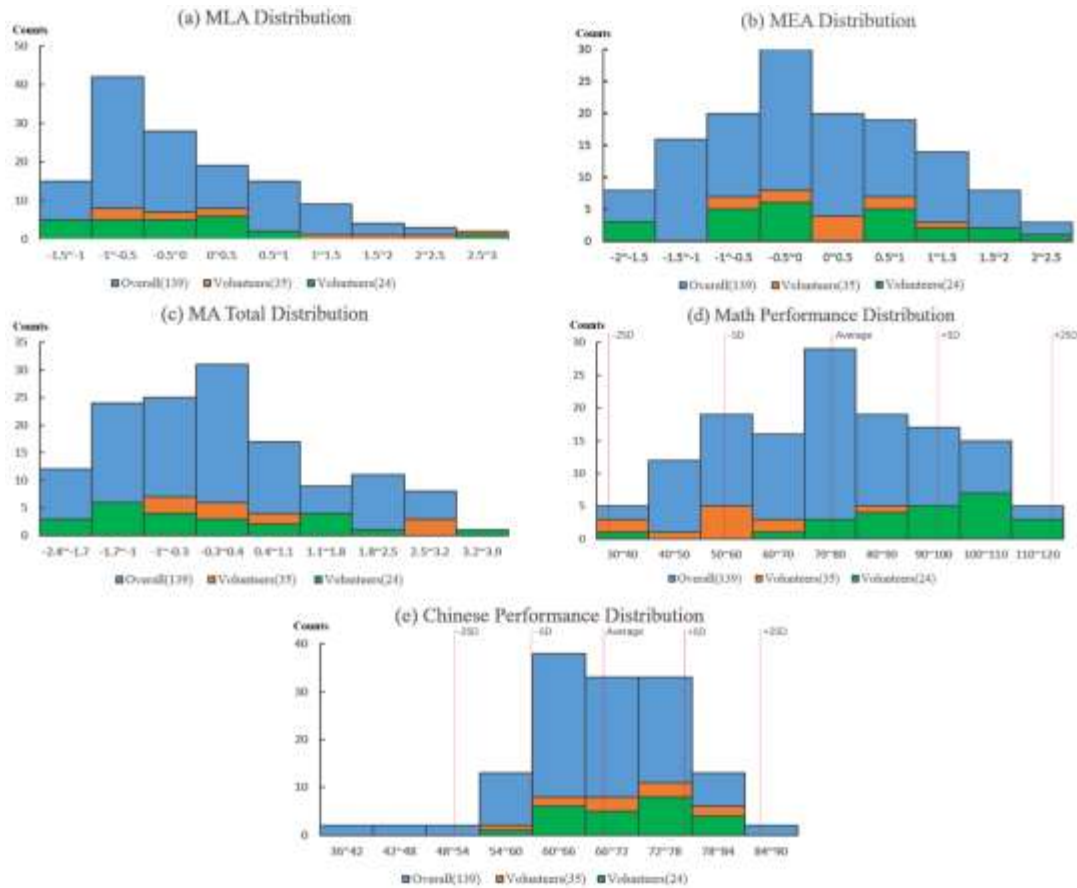


FIGURE 1. The distribution of the MA scores (MLA (a), MEA (b) and MA total (c)), academic performance (normalized Math scores (d) and normalized Chinese scores (e)) for all the 139 students, 35 volunteers who participated in the neurophysiological recording, and the 24 subjects who were included in the data analysis.

IV. DISCUSSION

In the present study, the neurophysiological signatures of MA were investigated by recording the SC and the HR in a cohort of high school students during their real finalterm math exam. The calculation of the pairwise correlation between the neurophysiological indicators and the MASC scores revealed that both SC and HR could reflect the individual's MEA, but in different exam periods. Specifically, SC was negatively correlated with MEA during the pre-exam period when the students were anticipating the exam, whereas HR was positively correlated with MEA at two later time windows during the exam. These

correlations remained significant after controlling for the students' math performance. By using these neurophysiological signatures in a regression model for predicting students' MEA scores, a correlation ($r = .394$, $p = .094$) was found between the predicted scores and their corresponding self-reported scores. Our findings, thus, provide neurophysiological evidence of MA in a real-life context and indicate the potential of implementing an objective measurement of trait MA using neurophysiological signals.

Unlike previous studies that often adopted a binary level approach dividing participants into high and low MA groups [16], [21],

[23], the present study adopted a correlational approach to investigate the neural signatures of trait MA. Indeed, scatterplots, as shown in Fig. 5, clearly show that the neurophysiological signatures were related to trait MA in a graded manner, which illustrate the importance of using the correlational approach. Since the correlational approach places greater emphasis on individual differences, it can better facilitate individual assessment of MA. With these MA-related

neurophysiologic signatures as independent variables, regression models with cross-validation indeed revealed a moderate positive correlation between the predicted and self-reported MEA scores ($r = .349$). Although the correlation was non-significant ($p = .094$), possibly due to the relatively small number of participants, our results demonstrated the possibility to measure individual's trait MA using neurophysiological signals.

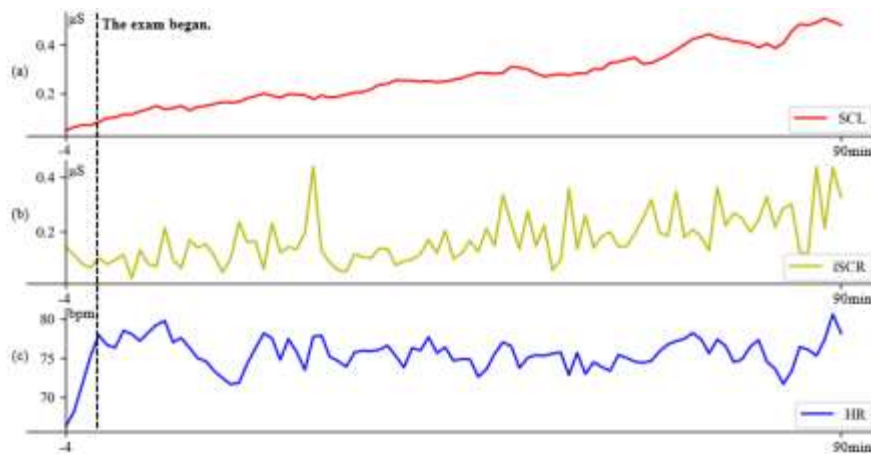


FIGURE 2. Temporal dynamics of the grand average SCL (a), iSCR (b) and HR (c).

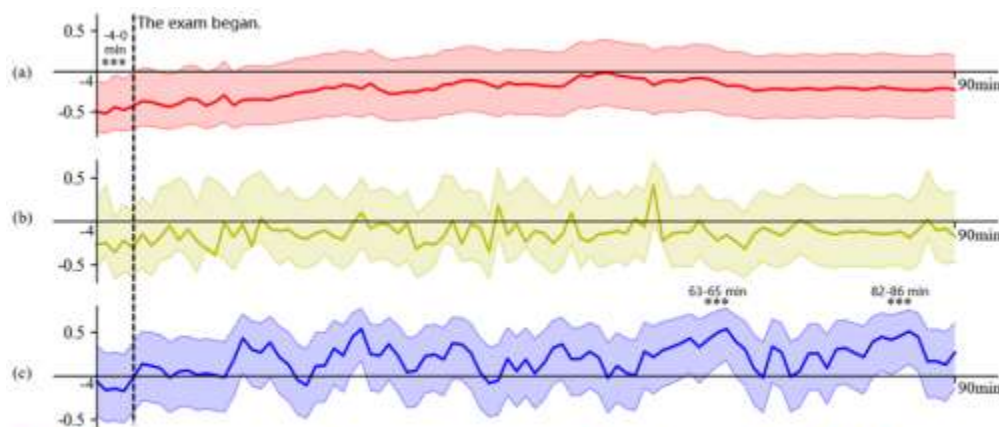


FIGURE 3. Temporal dynamic of the correlation values between MEA and SCL (a), iSCR (b), HR (c) during the experiment. The solid lines indicate the r values and the shadows indicate the 95% confidence intervals by bootstrapping. Significant clusters are indicated with black stars ***, $p < .001$.

Importantly Notably, distinct ANS signatures for anticipation and task periods

were found in the present study. In fact, previous CNS-based studies [18], [19], [21]–[23] have addressed possibly different neural processing in the anticipation and task periods. For instance, fMRI studies reported pain-related activity before math tasks and fear-related activity during math task for individuals with higher MA [18], [22]. EEG studies suggested that people with higher MA tended to use more attentional resources while expecting the arithmetic problems, and showed greater attentional bias toward arithmetic problems during the math task [21]. Although the functional roles of SCL

and HR have not been directly contrasted in MA, SCL and HR have been suggested to reflect independent neurophysiological mechanism in studies in other research fields: SCL is believed to reflect the engagement of attention [57] and consistent with arousal level [58], while HR reaction is more related to task or stimuli itself [24]. Regarding this study, our results provide ANS evidence in support of the differentiation of the two periods in the context of MA, at a 1-minute fine time scale. Details are discussed below.

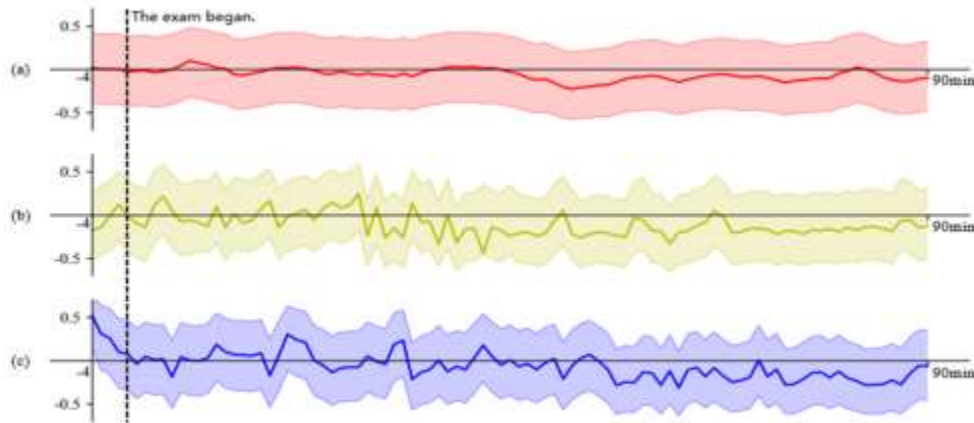


FIGURE 4. Temporal dynamic of the correlation values between MLA and (a) SCL, (b) ISCR, (c) HR during the experiment. The shadow indicates 95% confidence interval and the thicker lines indicate the average r values. Significant time clusters were indicated with black stars.

TABLE 2. The correlation between SCL (or HR) and MEA in the minutes of identified time clusters.

Linear Correlation		MEA (no control)		MEA (math performance controlled)	
		r	p	r	p
Cluster1	-4min	-.495*	.014	-.448*	.032
	-3	-.522**	.009	-.480*	.020
	-2	-.442*	.031	-.404	.056
	-1	-.478*	.018	-.441*	.035
	0	-.427*	.037	-.384	.070
Cluster2	63	.419*	.042	.343	.109
	64	.496*	.014	.469*	.024
	65	.541**	.006	.507*	.014
	82	.449*	.028	.528*	.010
Cluster3	83	.414*	.045	.413	.050
	84	.458*	.024	.482*	.020
	85	.513**	.010	.541**	.008
	86	.443*	.030	.453*	.030

Note. * $p < .05$, ** $p < .01$.

The positive correlations between HR and MEA in the exam period were consistent with many previous laboratory-based studies on MA or general anxiety. For example, when performing math tasks of increasing difficulty, students with higher MA were reported to have increased HR, while HR remained stable for the group with lower MA [16]. Regarding general anxiety, increased HR was reported in anxiety-induced tasks, such as stress interview [33] and public speech [24], [59]. In such studies, HR was regarded as an index of vigilance referring to the focus on external stimulus events [24], [60], [61]. In the present study, as the students received continuous stimuli of math problems during the exam, the students with higher MA might experience a higher level of vigilance, leading to increased HR for them.

Consequently, the correlations between HR and MEA remained positive during most of the exam period and significant correlations were found at the end of the exam. Our results provide validation of the laboratory-based findings on the relation between HR and MA in an actual exam situation.

More importantly, the 1-minute-based analysis provided more insights on the temporal dynamics of the cardiovascular responses to MA. Specifically, the positive

HR-MEA correlation reached its peak value after about one-hour exam time. As the students were supposed to be totally devoted to the exam items from the beginning of the exam, our non-significant HR-MEA correlations during the first ~60 minutes might suggest a possible suppression of MEA-related HR increases. The recruited high school students had extensive experiences in exams. To prepare for the National College Entrance Exam of China, they took exams every week and small quizzes every day. This suppression effect, therefore, could be explained by a coping mechanism developed by the students in order to reduce MA as reflected by the HR. Such a coping mechanism is not likely to be developed by participants in the laboratory-based tasks used in previous studies [16], [24], [33], in which the increase of the HR by MA or general anxiety was normally reported shortly after the onset of the tasks. Since the remaining time of the exam would have been announced in the classroom at the 60th minute and 80th minute, students with higher MEA were suggested to be more anxious than students with low MEA when dealing with the time pressure, resulting in a weakened suppression effect. The HR in students with higher MEA then increases shortly, possibly leading to a significant HR-MEA positive correlations at 63-65 minutes and 82-86 minutes.

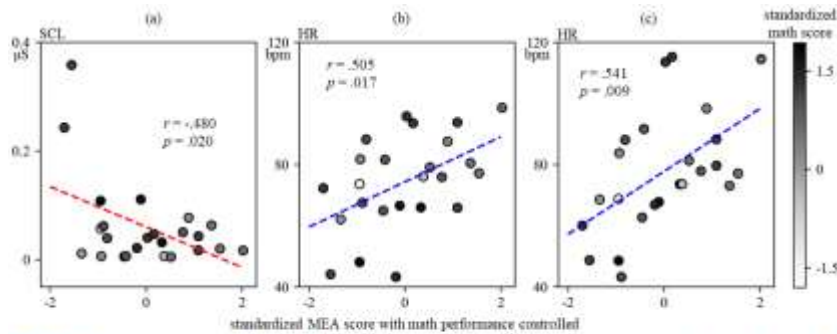


FIGURE 5. Scatterplots showing the relationship between neurophysiological signals and MEA with controlled math performance. X-axis refers to standardized MEA score with math performance controlled, and Y-axis refers to SCL or HR. Each data point represents a student and the darker the color, the higher the math score. Scatter plots of (a) SCL and MEA in the 4th minute at pre-exam stage, (b) HR and MEA in the 65th minute in the exam stage, and (c) HR and MEA in the 85th minute at the exam stage are presented respectively.

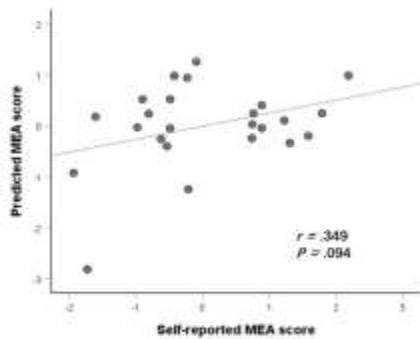


FIGURE 6. Scatterplot for the correlation between the predicted and self-reported MEA scores. Each dot represents the score from one participant (N = 24). The predicted score for each dot was obtained by using a leave-one-out cross-validation approach.

While further investigations are necessary to elucidate the underlying mechanism, our study reveals a complicated but practically useful result, as real-world MA inducing tasks are very likely to be associated with significant previous experiences.

The negative correlation between SCL and MEA in the anticipation (i.e. pre-exam) period, however, requires more efforts to interpret. Although the generally increasing trend of SCL throughout the recording time (Fig. 2a) is consistent with previous studies, for example, those about public speaking anxiety and anxiety towards threat of shock [24], [25], [62], people with a higher anxiety level is usually reported to be associated

with higher SCL [15], [24], [63], even at the anticipation period [64], [65]. The increased SCL during the anticipation period could be explained as an overall enhanced level of engagement for a more devoted preparation of the anxiety-related task [57], [61]. The seemingly inconsistent results of the negative correlation, might be due to the nature of the experimental paradigm in use. The real exam paradigm used in the present study is substantially different from these laboratory-based simulations, especially in terms of the potentially more severe consequence of failure. Accordingly, the students during the anticipation period could be heavily distracted by negative emotions other than anxiety, such as fear of failure, uncertainty about upcoming exam content, etc. [21]–[23]. These negative emotions would in general make the students less engaged, resulting in reduced SCL [57], [61], [66]. The students with higher MA could be more troubled by these negative emotions, leading to the observed negative correlation. Note that although the students might have developed their coping mechanism for math exams, they might not be able to adjust themselves before the actual onset of the exam (i.e. the anticipation

period), when the exam content is unknown. In support of our results, a recent study based on real-world settings also reported a negative correlation between the EDA mean and anxiety level during public speech, although no mechanistic interpretation was given [26].

As the first study to evaluate MA in an actual exam with neurophysiological recordings, there is room for further improvements. First, compared to the rich neurophysiological signals, the students' behavioral activities during the exam were very limited, making it difficult to further infer the mechanisms of the observed neurophysiological signatures. It would be preferable and feasible to obtain more behavioral data by camera recording, post-exam interviews, etc., for information such as when they actually finished the papers, how they coped with the time pressure, etc. Second, the neurophysiological indicators were defined based on their correlational significance, but a control baseline is missing. Although data from the non-significant time bins could serve to highlight the functional importance of these significant time points, ideally it is expected to have a baseline period beyond the anticipation and exam periods. Third, while the present study only used a simple correlational analysis to investigate the neurophysiological signature for MA, a deeper understanding of the underlying mechanisms of MA could benefit from the use of more advanced analytical methods, such as modeling the temporal characteristics of the neurophysiological signals by autoregression [67], exploring the spectral information of these

neurophysiological signals [68]–[70], etc. Last but not least, our findings might be limited by the relatively small sample size (due to feasibility concerns) and the biased distribution of the math performance of the volunteers towards the highperforming ends. Studies with a larger population and a more balanced sampling strategy are anticipated to provide further validation of the present findings.

ACKNOWLEDGMENT

The authors would like to thank the collaborating faculty members and students in the participant high school. They appreciate the excellent work in data preprocessing of our intern Mr. Z. Dong. (Zhilin Qu and Jingjing Chen contributed equally to this work.)

REFERENCES

- [1] G. Ramirez, S. T. Shaw, and E. A. Maloney, "Math anxiety: Past research, promising interventions, and a new interpretation framework," *Educ. Psychologist*, vol. 53, no. 3, pp. 145–164, Jul. 2018, doi: 10.1080/00461520.2018.1447384.
- [2] F. C. Richardson and R. M. Suinn, "The mathematics anxiety rating scale: Psychometric data," *J. Counseling Psychol.*, vol. 19, no. 6, pp. 551–554, Nov. 1972, doi: 10.1037/h0033456.
- [3] R. Hembree, "The nature, effects, and relief of mathematics anxiety," *J. Res. Math. Edu.*, vol. 21, no. 1, pp. 33–46, Jan. 1990, doi: 10.2307/749455.

[4] K. H. Dew and J. P. Galassi, "Mathematics anxiety: Some basic issues," *J. Counseling Psychol.*, vol. 30, no. 3, pp. 443–446, Jul. 1983, doi: 10.1037/0022-0167.30.3.443.

[5] J. Hunsley, "Cognitive processes in mathematics anxiety and test anxiety: The role of appraisals, internal dialogue, and attributions," *J. Educ. Psychol.*, vol. 79, no. 4, pp. 388–392, Dec. 1987, doi: 10.1037/0022-0663.79.4.388.

[6] OECD, "Does math make you anxious?" *PISA Focus*, OECD, Paris, France, Tech. Rep. 48, Feb. 2015, doi: 10.1787/5js6b2579tnx-en.

[7] M. H. Ashcraft, J. A. Krause, and D. R. Hopko, "Is math anxiety a mathematical learning disability?" in *Why is Math so Hard For Some Children?: The Nature and Origins of Mathematical Learning Difficulties and Disabilities*, D. Berch and M. Mazzocco, Eds. Baltimore, MD, USA: Brookes, Jan. 2017, pp. 329–348.

[8] H. Krinzinger, L. Kaufmann, and K. Willmes, "Math anxiety and math ability in early primary school years," *J. Psychoeducational Assessment*, vol. 27, no. 3, pp. 206–225, Jun. 2009, doi: 10.1177/0734282908330583.

[9] M. H. Ashcraft and J. A. Krause, "Working memory, math performance, and math anxiety," *Psychonomic Bull. Rev.*, vol. 14, no. 2, pp. 243–248, Apr. 2007, doi: 10.3758/BF03194059.

[10] G. Ramirez, E. A. Gunderson, S. C. Levine, and S. L. Beilock, "Math anxiety, working memory, and math achievement in

early elementary school," *J. Cognition Develop.*, vol. 14, no. 2, pp. 187–202, Apr. 2013, doi: 10.1080/15248372.2012.664593.

[11] R. M. Suinn, S. Taylor, and R. W. Edwards, "Suinn mathematics anxiety rating scale for elementary school students (MARS-E): Psychometric and normative data," *Educ. Psychol. Meas.*, vol. 48, no. 4, pp. 979–986, Dec. 1988, doi: 10.1177/0013164488484013.